

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ**

**Навчально-науковий інститут
Комп'ютерних інформаційних технологій**



**ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ
ЛІНГВІСТИЧНОГО АНАЛІЗУ
МІЖНАРОДНА
НАУКОВО-ТЕХНІЧНА КОНФЕРЕНЦІЯ
24–25 жовтня 2017 року**

Тези доповідей

Київ 2017

ЗМІСТ

О.П. Нечипорук НАЛАШТУВАННЯ ЕКСПЕРТНОЇ МОДЕЛІ ДІАГНОСТУВАННЯ БАГАТОРІВНЕВОГО ТЕХНІЧНОГО ОБ'ЄКТА.....	6
Я.А. Белозьорова ІДЕНТИФІКАЦІЯ ДИКТОРА НА ОСНОВІ ПОЛІПШЕНОГО АЛГОРИТМУ ВИДІЛЕННЯ ЧАСТОТИ ОСНОВНОГО ТОНУ	9
Т.І. Конрад ДОСЛІДЖЕННЯ АРХІТЕКТУРИ МУЛЬТИМОДАЛЬНИХ ІНТЕЛЕКТУАЛЬНИХ ТРАНСПОРТНИХ ІНФОРМАЦІЙНИХ СИСТЕМ	10
Д.В. Ландэ, Б.А. Березин, О.Ю. Павленко РАЗРАБОТКА АЛГОРИТМА СЕГМЕНТАЦИИ СЛОВ ДЛЯ СИСТЕМ МОНИТОРИНГА НАЦИОНАЛЬНЫХ ИНТЕРНЕТ- РЕСУРСОВ.....	11
Є.Б. Артамонов АДАПТИВНІ НАВЧАЛЬНІ РЕСУРСИ ПРИ ВИКОРИСТАННІ СЦЕНАРНОГО ПІДХОДУ В СИСТЕМАХ ЕЛЕКТРОННОГО НАВЧАННЯ.....	12
О.Л. Ляхов, Н.А. Фурсова АЛГОРИТМ КЛАСТЕРИЗАЦІЇ ДЛЯ DATA MINING У СОЦІАЛЬНО-ЕКОНОМІЧНИХ ДОСЛІДЖЕННЯХ	13
А.І. Вавіленкова ПРОЕКТ ТЕХНОЛОГІЇ АВТОМАТИЧНОГО ЛІНГВІСТИЧНОГО АНАЛІЗУ ТЕКСТОВИХ ДОКУМЕНТІВ	15
І.В. Вантух ТЕХНОЛОГІЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ТА ОБРОБКИ ТЕКСТІВ TEXT MINING.....	16
І.В. Огірко, О.І. Огірко ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА МОДЕЛІ ЛІНГВІСТИЧНОГО АНАЛІЗУ ТЕКСТУ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ОЦІНЮВАННЯ ЗНАНЬ	17
О.В. Кравченко СИСТЕМИ ПОРІВНЯЛЬНОГО АНАЛІЗУ ЕЛЕКТРОННИХ ТЕКСТІВ	18

*О.Л. Ляхов, д.т.н., Н.А. Фурсова
Полтавський національний технічний університет
імені Юрія Кондратюка*

АЛГОРИТМ КЛАСТЕРИЗАЦІЇ ДЛЯ DATA MINING У СОЦІАЛЬНО-ЕКОНОМІЧНИХ ДОСЛІДЖЕННЯХ

У кластерному аналізі використовується поняття «схожості» об'єктів, що потребує подальшої формалізації фактично у кожному класі задач, що розв'язуються. Наслідком, на наш погляд, є неоднозначність вибору алгоритму із величезної кількості і неприродні, на наш погляд, параметри алгоритмів (наперед визначена кількість кластерів і т. і.). Ці функції реально виконуються людиною-експертом, що буде джерелом труднощів при подальшій автоматизації розв'язування задач кластеризації.

У роботі [1] пропонується розглядати кластер, як клас, утворений деяким відношенням феноменологічної симетрії, і «схожість» замінити відомим, строго формалізованим та узагальненим поняттям «близькості». Вважається, що такий підхід веде до зменшення кількості застосовуваних моделей алгоритмів, можливості єдиного підходу до кластеризації об'єктів з числовими і нечисловими характеристиками. Перспективною, на наш погляд, виглядає і можливість використання такого підходу у лінгвістичному аналізі на основі теоретико-множинних моделей мов.

Проте, на сьогодні ставляться більш скромні задачі апробації підходу та дослідження ефективності алгоритмів. У роботі [1] наведені результати розв'язування простих тестових прикладів. У даній роботі наведені результати апробації підходу з використанням загальновідомої метричної інтерпретації «близькості» у реальних соціально-економічних дослідженнях.

На підставі офіційних статистичних даних були досліджені закономірності розподілу та темпи зміни середньої заробітної плати за усіма регіонами України на протязі 2000-2014 років.

Кластери, знайдені у розподілі даних за регіонами, відповідають відомим результатам [2], що разом із [1] свідчить про адекватність запропонованого підходу та його алгоритмізації на досліджуваних задачах.

Кластеризація за темпами зміни даних дала змогу виявити нові, приховані закономірності, які з експертної точки зору можуть свідчити про можливу доцільність зміни пріоритетів у видах економічної діяльності регіонів, порівняно із загально визнаними. Наприклад, для промислової Миколаївської області – це транспортні послуги, для сільськогосподарської Житомирської області – це страховий бізнес та інші.

Використані джерела

1. *Ляхов О.Л.* Алгоритм кластеризації для Data Mining / *О.Л. Ляхов, Н.А. Фурсова* // Математичне та імітаційне моделювання систем. МОДС 2017 : тези доп. Дванадцятої міжнар. наук.-практ. конф. (м. Чернігів, 26–29 черв. 2017 р.). – Чернігів : ЧНТУ, 2017. – С. 425-428.

2. *Регіональний людський розвиток: Статистичний бюлетень* / Відп. за вип. О.О. Кармазіна. – К. : Державна служба статистики України, 2017. – 65 с.